



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

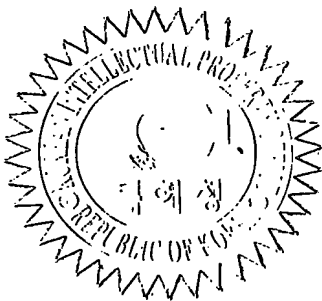
PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

출원번호 : 10-2003-0077328
Application Number

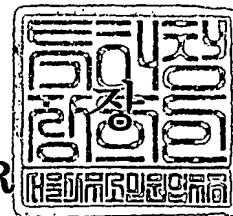
출원년월일 : 2003년 11월 03일
Date of Application NOV 03, 2003

출원인 : 삼성전자주식회사 외 1명
Applicant(s) SAMSUNG ELECTRONICS CO., LTD., et al.



2005 년 01 월 26 일

특 허 청
COMMISSIONER



BEST AVAILABLE COPY

【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0009
【제출일자】	2003.11.03
【국제특허분류】	H04N
【발명의 명칭】	게이즈 디텍션을 이용한 비디오 데이터 처리 장치 및 방법
【발명의 영문명칭】	Apparatus and method for processing video data using gaze detection
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【출원인】	
【명칭】	학교법인 고황재단
【출원인코드】	2-1999-901357-1
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2003-003435-0
【포괄위임등록번호】	2003-074840-7
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2003-003436-7
【포괄위임등록번호】	2003-074841-4
【발명자】	
【성명의 국문표기】	박광훈
【성명의 영문표기】	PARK, Gwang Hoon
【주민등록번호】	630812-1149111
【우편번호】	463-030
【주소】	경기도 성남시 분당구 분당동(셋별마을) 414동 804호
【국적】	KR

【취지】

특허법 제42조의 규정에 의하여 위와 같이 출원합니다. 대리인
이영필 (인) 대리인
이해영 (인)

【수수료】

【기본출원료】 20 면 29,000 원

【가산출원료】 14 면 14,000 원

【우선권주장료】 0 건 0 원

【심사청구료】 0 항 0 원

【합계】 43,000 원

【첨부서류】

1. 요약서·명세서(도면)_1통

【요약서】**【요약】**

게이즈 디텍션을 이용한 비디오 데이터 처리 장치 및 방법이 개시된다. 본 발명에 따른 비디오 데이터 처리 장치는, 입력 비트스트림을 스케일러블 디코딩하는 스케일러블 디코더; 게이즈 디텍션을 이용하여 디스플레이 되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하여, 상기 관심 영역의 위치 정보를 출력하는 관심영역 결정부; 및 상기 관심영역 결정부로부터 입력받은 상기 관심 영역의 위치 정보에 따라, 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하고, 상기 선택된 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하도록 상기 스케일러블 디코더를 제어하는 제어부를 포함한다. 본 발명에 따르면, 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지하여 관심 영역을 스케일러블 디코딩하여 그 화질을 향상시킴으로써 디코더의 로드를 감소시키거나 데이터 통신 채널의 대역 폭의 한계를 극복할 수 있다.

【대표도】

도 4

【명세서】

【발명의 명칭】

게이즈 디텍션을 이용한 비디오 데이터 처리 장치 및 방법{Apparatus and method for processing video data using gaze detection}

【도면의 간단한 설명】

도 1은 다수의 VOP로 분할된 영상 프레임을 나타내는 도면,

도 2는 MPEG-4 인코더의 블록도의 일 예를 나타내는 도면,

도 3은 MPEG-4 디코더의 블록도의 일 예를 나타내는 도면,

도 4는 본 발명의 제1 실시예에 따른 비디오 데이터 처리 장치의 블록도를 나타내는 도면,

도 5는 도 4에 도시된 관심영역 결정부의 블록도의 일 예를 나타내는 도면,

도 6a 및 도 6b는 게이즈 디텍션 방법의 일 예를 설명하기 위한 도면,

도 7은 도 4에 도시된 디코더의 블록도의 일 예를 나타내는 도면,

도 8은 입력 비트스트림에서 영상 객체별 비트스트림을 추출하는 과정을 설명하기 위한 도면,

도 9는 서브 스케일러블 디코더의 블록도의 일 예를 나타내는 도면,

도 10a 및 도 10b는 콘텐츠별로 스케일러블 코딩 및 디코딩 경우, 본 발명에 따라 관심 있는 콘텐츠의 화질이 개선되었음을 나타내는 도면,

도 11a 및 도 11b는 프레임별로 스케일러블 코딩 및 디코딩 경우, 본 발명에 따라 관심 있는 프레임의 화질이 개선되었음을 나타내는 도면,

도 12는 본 발명의 제2 실시예에 따른 비디오 데이터 처리 장치의 블록도이다.

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

- <13> 본 발명은 비디오 데이터 처리 장치 및 방법에 관한 것으로, 보다 상세하게는 게이즈 디텍션(Gaze detection)을 이용하여 디스플레이 되는 영상 중에서 사용자가 특히 관심 있는 영역의 화질을 개선시킬 수 있는 비디오 데이터 처리 장치 및 방법에 관한 것이다.
- <14> 과거의 비디오 데이터 부호화 기술은 단지 비디오 데이터를 압축하여 저장하거나 전달하는 것에 초점이 맞추어져 있었으나 현재의 기술은 비디오 데이터의 상호 교환과 사용자 조작성(User Interaction) 부여에 초점이 맞추어져 있다.
- <15> 예컨대, 비디오 압축 기술의 국제표준 중 하나인 MPEG-4 Part 2의 비디오 압축 기술은 영상 프레임 내에 포함된 콘텐츠(Contents)별로 코딩해서 전송하는 VOP(Video Object Plane)단위 코딩 기법을 채택하고 있다. 도 1은 MPEG-4 비디오 부호화 표준에 따라 다수의 VOP로 분할된 영상 프레임을 나타내는 도면이다. 도 1을 참조하면, 영상 프레임(1)은 배경 영상에 해당하는 VOP 0(11)와 프레임(1) 내에 포함된 각각의 콘텐츠 별로 VOP 1(13), VOP 2(15), VOP 3(17) 및 VOP 4(19)로 분할된다.
- <16> 도 2는 MPEG-4 인코더의 블록도이다. 도 2를 참조하면, MPEG-4 인코더는 입력 영상을 VOP 단위로 분할하여 각 VOP를 출력하는 VOP 정의부(21), 각각의 VOP를 인코딩하는 다수의 VOP 인코더(23,25,27) 및 인코딩된 VOP별 데이터를 다중화하여 비트스트림을 생성하는 다중화부

(MUX)(29)를 포함한다. VOP 정의부(21)는 각 콘텐츠의 모양정보(Shape Information)을 이용하여 영상 프레임 내의 콘텐츠 단위로 VOP를 정의한다.

<17> 도 3은 MPEG-4 디코더의 블록도이다. 도 3을 참조하면, MPEG-4 디코더는 입력 비트스트림으로부터 각 VOP별 비트스트림을 선택하여 역다중화하는 역다중화부(31), 각 VOP별 비트스트림을 디코딩하는 다수의 VOP 디코더(33,35,37) 및 VOP 합성부(39)를 포함한다.

<18> 전술한 바와 같이, MPEG-4는 VOP 단위로 영상을 인코딩 및 디코딩하므로 콘텐츠 기반의 조작성을 사용자에게 부여할 수 있다.

<19> 한편, 일반적으로 영상 데이터는 MPEG과 같은 데이터 압축 표준을 지원하는 인코더에 의해 인코딩된 후 비트스트림의 형태로 정보저장매체에 저장되거나 통신 채널을 통해 전송된다. 하나의 비트스트림으로부터 공간 해상도가 서로 다른 영상을 재생할 수 있거나 시간당 재생 프레임의 개수, 즉 시간 해상도가 다른 영상을 재생할 수 있다면, 그 비트스트림은 스케일러블(scalable)하다고 한다. 전자가 비트스트림이 공간적으로 스케일러블(spatial scalable)한 경우이고, 후자는 시간적으로 스케일러블(temporal scalable)한 경우이다.

<20> 스케일러블한 비트스트림은 베이스 레이어(base layer) 데이터와 인핸스먼트 레이어(enhancement layer) 데이터를 포함한다. 공간적으로 스케일러블한 비트스트림의 응용예를 설명하면, 디코더는 베이스 레이어 데이터를 디코딩함으로써 일반 텔레비전급 정도의 화질을 가지는 영상을 재생할 수 있으나 베이스 레이어 데이터를 이용하여 인핸스먼트 레이어 데이터까지 디코딩하는 경우에는 HD TV(High Definition TV)급 영상을 재생할 수 있다.

<21> MPEG-4 또한 스케일러빌리티 기능을 지원한다. 즉, 각 VOP 단위로 스케일러블 인코딩을 수행할 수 있어, VOP 단위로 시간적 또는 공간적 해상도가 다른 영상을 재생할 수 있다.

<22> 한편, 초대형 화면의 영상 또는 다수의 프레임 영상으로 이루어진 멀티 프레임 영상을 종래 기술에 따라 인코딩 시 전송할 비디오 데이터 양이 매우 많아진다. 나아가 영상을 스케일 러블 코딩한 경우에는 전송할 비디오 데이터 양은 더 많아지므로 데이터 전송 채널의 대역 폭 (bandwidth)의 제한 또는 디코더의 성능의 한계로 인해 높은 화질의 영상을 재생하여 사용자에게 보여주는데 어려움이 있다.

【발명이 이루고자 하는 기술적 과제】

<23> 따라서 본 발명이 이루고자 하는 기술적 과제는, 데이터 전송 채널의 대역 폭 (bandwidth)의 제한 또는 디코더의 성능의 한계가 존재하는 상황에서 사용자에게 디스플레이 되는 영상에서 사용자가 응시하는 관심 영역의 영상의 화질을 향상시킬 수 있는 비디오 데이터 처리 방법을 제공하는데 있다.

<24> 본 발명이 이루고자 하는 다른 기술적 과제는, 데이터 전송 채널의 대역 폭(bandwidth)의 제한 또는 디코더의 성능의 한계가 존재하는 상황에서 사용자에게 디스플레이 되는 영상에서 사용자가 응시하는 관심 영역의 영상의 화질을 향상시킬 수 있는 비디오 데이터 처리 장치를 제공하는데 있다.

【발명의 구성 및 작용】

<25> 상기 과제를 이루기 위해, 본 발명의 일 측면에 의한 비디오 데이터 처리 방법은,

<26> 게이즈 디텍션을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계; 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하는 단계; 및 상기 영상 객체

의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하는 단계를 포함한다.

<27> 상기 과제를 이루기 위해, 본 발명의 다른 측면에 의한 비디오 데이터 처리 방법은,

<28> 상대방 장치로부터 수신한 이전 비트스트림을 디코딩하여 사용자에게 디스플레이하는 단계; 게이즈 디택션을 이용하여 디스플레이되는 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계; 상기 관심 영역의 위치 정보를 상기 상대방 장치로 송신하는 단계; 상기 상대방 장치로부터 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 포함하는 현재 비트스트림을 수신하는 단계; 및 상기 현재 비트스트림을 스케일러블 디코딩하는 단계를 포함한다.

<29> 상기 다른 과제를 이루기 위해, 본 발명의 일 측면에 의한 비디오 데이터 처리 장치는,

<30> 입력 비트스트림을 스케일러블 디코딩하는 스케일러블 디코더; 게이즈 디택션을 이용하여 디스플레이 되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하여, 상기 관심 영역의 위치 정보를 출력하는 관심영역 결정부; 및 상기 관심영역 결정부로부터 입력받은 상기 관심 영역의 위치 정보에 따라, 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하고, 상기 선택된 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하도록 상기 스케일러블 디코더를 제어하는 제어부를 포함한다.

31> 상기 다른 과제를 이루기 위해, 본 발명의 다른 측면에 의한 비디오 데이터 처리 장치는,

<32> 입력 비트스트림을 스케일러블 디코딩하는 스케일러블 디코더; 상대방 장치로부터 수신하여 디코딩된 후 사용자에게 디스플레이 되는 영상에서 상기 사용자가 응시하는 관심 영역의 위치를 게이즈 디택션을 이용하여 결정하여, 상기 관심 영역의 위치 정보를 출력하는 관심영역 결정부; 및 상기 관심 영역의 위치 정보를 상기 상대방 장치로 송신하는 데이터 통신부를 포함하고, 상기 스케일러블 디코더는 상기 상대방 장치로부터 수신한, 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 포함하는 현재 비트스트림을 스케일러블 디코딩하는 것을 특징으로 한다.

<33> 이하, 첨부된 도면들을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명한다.

<34> 본 발명은 게이즈 디택션 방법을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지하여 관심 영역의 화질을 스케일러블 디코딩을 수행하여 향상시킨다. 본 발명은 특히, 공간적 해상도가 큰 대형 화면의 영상, 예컨대 네 면의 벽면에 설치된 대형 디스플레이 장치에 의해 표시되는 영상 또는 다수의 프레임 영상으로 이루어진 멀티 프레임 영상을 사용자에게 디스플레이 하는 경우 유용하다. 왜냐하면 공간적 해상도가 매우 큰 영상을 스케일러블 코딩한 경우, 전송할 비디오 데이터 양이 매우 많으므로 데이터 전송 채널의 대역 폭의 제한 또는 디코더의 성능의 한계로 인해 높은 화질의 영상을 재생하여 사용자에게 보여주는데 어려움이 있기 때문이다.

<35> 게이지 디택션 방법을 이용하여 탐지된 관심 영역의 화질을 스케일러블 디코딩을 수행하여 향상시키기 위해 본 발명은 다음의 두 가지 실시예를 개시한다. 제1 실시예에 따르면, 게이즈 디택션 방법을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지한 후, 관심 영역을 포함하는 영상 객체만을 스케일러블 디코딩을 수행하여 관심 영

역의 화질을 향상시키고 다른 영상 객체들은 베이스 레이어 데이터만 디코딩한다. 즉, 스케일러블 디코더의 성능의 한계를 고려하여 관심 영역의 화질을 개선하기 위한 실시예이다.

<36> 제2 실시예에 따르면, 게이즈 디택션 방법을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지한 후, 본 발명에 따른 비디오 데이터 처리 장치로 비트스트림을 전송하는 상대방 장치(encoder)로, 탐지된 관심 영역의 위치 정보를 전송한다. 탐지된 관심 영역의 위치 정보를 수신한 그 상대방 장치는 관심 영역을 포함하는 영상 객체만을 스케일러블 인코딩하고 다른 영상 객체들은 베이스 레이어 인코딩만 수행함으로써 통신 채널을 통해 전송할 데이터 양을 크게 줄인다. 즉, 데이터 통신 채널의 대역 폭의 한계를 고려하여 관심 영역의 화질을 개선하기 위한 실시예이다.

<37> 데이터 통신 채널로는 PSTN, ISDN, Internet, ATM Network 또는 무선 통신망(Wireless communication network) 등과 같은 다양한 전송 매체가 이용될 수 있다.

<38> 영상 객체(Video object)는 멀티프레임 영상인 경우, 하나의 프레임 또는 MPEG-4와 같이 하나의 프레임 영상을 그 프레임 영상에 포함된 영상 콘텐츠별로 분리하여 코딩한 경우, 각 영상 콘텐츠(i.e VOP)를 말한다.

<39> 이하, 전술한 본 발명의 두 가지 실시예를 첨부한 도면을 참조하여 보다 상세하게 설명한다.

<40> I. 제1 실시예

<41> 도 4는 본 발명의 제1 실시예에 따른 비디오 데이터 처리 장치의 블록도이다. 도 4를 참조하면, 비디오 데이터 처리 장치는 관심영역 결정부(110), 제어부(130) 및 디코더(150)를 포함한다.

- <42> 관심영역 결정부(110)는 게이지 디텍션을 이용하여 디스플레이 장치(도시되지 않음)를 통해 사용자에게 디스플레이 되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하여 관심 영역의 위치 정보를 제어부(130)로 출력한다.
- <43> 제어부(130)는 관심영역 결정부(110)로부터 입력받은 관심 영역의 위치 정보에 따라, 입력 비트스트림에서 그 관심 영역을 포함하는 영상 객체(Video object)의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하고, 선택된 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하도록 디코더(150)를 제어한다.
- <44> 디코더(150)는 제어부(130)의 제어에 따라 입력 비트스트림에 대해 스케일러블 디코딩을 수행하는 스케일러블 디코더이다. 디코더(150)는 제어부(130)의 제어에 따라, 입력 비트스트림에서 사용자가 응시하는 관심 영역을 포함하는 영상 객체의 인핸스먼트 레이어 비트스트림을 선택하여 스케일러블 디코딩을 수행함으로써 관심 영역의 화질을 향상시킨다. 또한, 제어부(130)의 제어에 따라 관심 영역을 포함하는 영상 객체가 아닌 다른 영상 객체의 인핸스먼트 레이어 비트스트림은 디코딩하지 않고 베이스 레이어 데이터만을 디코딩함으로써 디코더(150)에 걸리는 로드(load)를 감소시킨다.
- <45> 도 5는 도 4에 도시된 관심영역 결정부(110)의 블록도의 일 예를 나타내는 도면이다. 도 5를 참조하면 관심영역 결정부(110)는 머리 부분을 중심으로 사용자를 촬영하는 비디오 카메라(111) 및 비디오 카메라(111)로부터 입력된 사용자의 동영상을 분석하여 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 게이지 디텍션부(113)를 포함한다.
- <46> 게이지 디텍션은 사용자의 머리의 움직임 및/또는 눈동자의 움직임을 추정하여 사용자가 응시하는 위치를 탐지하는 방법으로서 구체적인 실시예는 매우 다양하다. 예컨대, 한국 공개특허 공보 제2000-0056563호에는 게이지 디텍션 방법의 일 예가 기재되어 있다.

<47> 도 6a 및 도 6b는 전술한 한국 공개 특허 공보에 기재된 게이지 디텍션 방법을 설명하기 위한 도면이다. 사용자는 디스플레이 장치, 예컨대 모니터에 디스플레이 되는 정보를 주로 눈동자를 움직여서 특정 부분의 정보를 인지하거나, 머리를 움직여서 인지하게 된다. 이러한 점에 착안하여 모니터 위 또는 사용자의 머리 부분을 촬영하기에 적합한 위치에 설치된 비디오 카메라를 통하여 촬영된 사용자에 대한 영상 정보를 분석하여 사용자가 응시하는 모니터 상의 위치를 탐지한다.

<48> 도 6a는 사용자가 정면으로 모니터를 응시할 때의 사용자의 두 눈, 코 및 입의 위치를 나타낸다. 점 P1 및 P2는 두 눈의 위치, P3는 코의 위치를 나타내고 P4 및 P5는 입의 양 끝 위치를 나타낸다.

<49> 도 6b는 사용자가 머리를 움직여 모니터의 정면이 아닌 다른 곳을 응시할 때의 사용자의 두 눈, 코 및 입의 위치를 나타낸다. 마찬가지로 점 P1 및 P2는 두 눈의 위치, P3는 코의 위치를 나타내고 P4 및 P5는 입의 양 끝 위치를 나타낸다. 따라서 게이지 디텍션부(113)는 다섯가지 종류의 위치의 변화를 감지하여 사용자가 응시하는 모니터 상의 위치를 탐지할 수 있다.

<50> 본 발명에 따른 게이지 디텍션 방법은 전술한 실시예에 한정하지 않으며 어떠한 게이지 디텍션 방법이라도 무방하다. 또한, 본 발명에 따른 관심영역 결정부(110)는 다양한 형태로 구현할 수 있다. 예컨대, 사용자의 모습을 촬영할 수 있는 소형의 카메라 및 사용자의 머리의 움직임을 감지할 수 있는 장치가 장착된 헬멧, 고글(goggle) 또는 안경의 형태로 제작할 수 있다. 사용자는 게이지 디텍션 기능을 구비한 헬멧과 같은 형태의 특수 장비를 착용하고, 그 특수 장비는 사용자가 응시하는 관심 영역의 위치를 탐지한 후 제어부(130)로 탐지된 관심 영역의 위치 정보를 유선 또는 무선 통신 방식을 이용하여 전달한다. 게이지 디텍션 기능을 구비한 헬멧과 같은 형태의 특수 장비는 이미 상용화되어 있다. 예컨대, 전투용 헬리콥터의 조종사

들이 게이지 디텍션 기능을 구비한 헬멧을 착용하여 기관총의 사격을 위한 조준에 사용하고 있다.

<51> 도 7은 도 4에 도시된 디코더(150)의 블록도의 일 예를 나타내는 도면이다. 도 7을 참조하면, 디코더(150)는 시스템 역다중화부(151), 영상 객체 역다중화부(153) 및 스케일러블 디코더(155)를 포함한다. 스케일러블 디코더(155)는 영상 객체 단위로 각각 스케일러블 디코딩을 수행하는 다수의 서브 스케일러블 디코더(155a, 155b, 155c)를 포함한다.

<52> 시스템 역다중화부(151)는 입력 비트스트림을 시스템 비트스트림, 비디오 스트림 및 오디오 스트림 등으로 역다중화하여 출력한다. 특히, 시스템 역다중화부(151)는 제어부(130)의 제어에 따라 입력 비트스트림에서 사용자가 응시하는 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림과, 관심 영역을 포함하지 않는 나머지 영상 객체들의 경우에는 베이스 레이어 비트스트림만을 선택하여 영상 객체 역다중화부(153)로 출력한다. 즉, 관심 영역을 포함하지 않는 나머지 영상 객체들의 인헨스먼트 레이어 비트스트림은 디코딩되지 않도록 하기 위해 영상 객체 역다중화부(153)로 출력하지 않는다.

<53> 도 8은 입력 비트스트림에서 영상 객체별 비트스트림을 추출하는 과정을 설명하기 위한 도면이다. 입력 비트스트림이 MPEG-4 part 2 규격에 따라 생성된 비트스트림인 경우, 입력 비트스트림은 Scene Description Stream(210) 및 Object Description Stream(230)과 같은 시스템 비트스트림을 포함한다. Scene Description Stream(210)은 하나의 영상의 구조를 설명하는 Interactive Scene Description(220)을 포함하는 비트스트림으로서 Interactive Scene Description(220)은 트리 구조(220)를 가진다.

- <54> Interactive Scene Description(220)은 하나의 영상(300)에 포함된 VOP 0(270), VOP 1(280) 및 VOP 2(290)의 위치 정보, 각 VOP의 오디오 데이터 정보 및 비디오 데이터 정보를 포함한다.
- <55> Object Description Stream(230)은 각 VOP의 오디오 비트스트림의 위치 정보 및 비디오 비트스트림의 위치 정보를 포함한다. 도 8을 참조하면, 사용자가 응시하는 관심 영역을 포함하는 영상 객체, 즉 VOP는 VOP 0(270)이다. 시스템 역다중화부(151)는 제어부(130)의 제어에 따라 관심영역 결정부(110)로부터 입력받은 관심 영역의 위치정보와, 입력 비트스트림에 포함된 Scene Description Stream(210) 및 Object Description Stream(230)에 포함된 정보를 비교하여 입력 비트스트림에서 사용자가 응시하는 VOP 0(270)의 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림을 포함하는 비주얼 스트림(240)과, 관심 영역을 포함하지 않는 나머지 영상 객체들의 경우에는 베이스 레이어 비트스트림(250,260)만을 선택/추출하여 영상 객체 역다중화부(153)로 출력한다.
- <56> 영상 객체 역다중화부(153)는 비트스트림에 포함된 각각의 영상 객체의 비트스트림을 역다중화하여 각 영상 객체의 비트스트림을 스케일러블 디코더(155)의 대응하는 각 서브 스케일러블 디코더(155a,155b,155c)로 출력한다.
- <57> 만약, 영상객체 0가 관심 영역을 포함하는 영상 객체라면 서브 스케일러블 디코더 0(155a)로는 영상 객체 0의 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림이 입력되어 서브 스케일러블 디코더 0(155a)는 스케일러블 디코딩을 수행할 것이다. 따라서 영상 객체 0는 고품질의 영상이 재생된다. 그 외 다른 서브 스케일러블 디코더(155b,155c)로는 각 영상 객체의 베이스 레이어 비트스트림만 입력되어 베이스 레이어 디코딩만 수행되어 저화질의 영상이 재생된다.

- <58> 도 9는 서브 스케일러블 디코더의 블록도의 일 예를 나타낸다. 도 9를 참조하면, 서브 스케일러블 디코더는 인헨스먼트 레이어 디코더(410), 중간 처리부(Mid-processor)(430), 베이스 레이어 디코더(450) 및 후처리부(Post-Processor)(470)를 포함한다.
- <59> 베이스 레이어 디코더(450)는 베이스 레이어 비트스트림을 입력받아 베이스 레이어 디코딩을 수행한다. 인헨스먼트 레이어 디코더(410)는 인헨스먼트 레이어 비트스트림과 중간 처리부(Mid-processor)(430)로부터 입력된 베이스 레이어 비트스트림을 이용하여 인헨스먼트 레이어 디코딩을 수행한다. 중간 처리부(Mid-processor)(430)는 베이스 레이어 비트스트림이 인코더에 의해 공간적 스케일러블 인코딩된 비트스트림인 경우, 베이스 레이어 디코딩된 베이스 레이어 데이터를 업 샘플링하여 공간적 해상도를 증가시킨 후 인헨스먼트 레이어 디코더(410)로 제공한다. 후처리부(Post-Processor)(470)는 베이스 레이어 디코더(450)와 인헨스먼트 레이어 디코더(410)로부터 각각 디코딩된 베이스 레이어 데이터 및 인헨스먼트 레이어 데이터를 입력받아 두 데이터를 결합한 후 스무딩(smoothing) 등의 신호 처리를 한다.
- <60> 도 10a 및 도 10b는 콘텐츠별로 스케일러블 코딩 및 디코딩 경우, 본 발명에 따라 관심 있는 콘텐츠의 화질이 개선되었음을 나타내는 도면이다. 도 10a 및 도 10b를 참조하면, 영상은 다수의 콘텐츠(13,15,17,18)를 포함한다.
- <61> 도 10a는 종래 기술에 따라 재생한, 다수의 콘텐츠(13,15,17,18)를 포함한 영상을 나타낸다. 종래 기술에 따르면 데이터 전송 채널의 대역 폭의 제한 또는 디코더의 성능의 한계로 인해 스케일러블 비트스트림을 전송하지 못하거나 스케일러블 비트스트림을 디코더가 수신하더라도 디코더의 성능의 제약으로 인해 선명하지 못한 영상을 재생한다.
- <62> 도 10b는 본 발명에 따라, 사용자가 응시하는 관심 영역의 화질이 개선된 재생 영상을 나타낸다. 본 발명에 따르면, 게이즈 디텍션 방법을 이용하여 디스플레이되는 현재 영상에서

사용자가 응시하는 관심 영역의 위치를 탐지한 후, 관심 영역을 포함하는 영상 객체(13)만을 스케일러블 디코딩을 수행하여 관심 영역의 화질을 향상시키고 다른 영상 객체들(15,17,18)은 베이스 레이어 데이터만 디코딩한다.

63> 도 11a 및 도 11b는 멀티 프레임 영상인 경우 프레임별로 스케일러블 코딩 및 디코딩 경우, 본 발명에 따라 관심있는 프레임의 화질이 개선되었음을 나타내는 도면이다. 도 11a 및 도 11b를 참조하면, 디스플레이 장치(500)를 통해 다수의 프레임 영상(510,530)을 포함하는 멀티 프레임 영상이 디스플레이 된다.

64> 도 11a는 종래 기술에 따라 재생한, 프레임 영상(510,530)을 포함하는 멀티 프레임 영상을 나타낸다. 데이터 전송 채널의 대역 폭의 제한 또는 디코더의 성능의 한계로 인해 스케일러블 비트스트림을 전송하지 못하거나 스케일러블 비트스트림을 디코더가 수신하더라도 디코더의 성능의 제약으로 인해 선명하지 못한 멀티 프레임 영상을 재생한다.

65> 도 11b는 본 발명에 따라, 사용자가 응시하는 관심 영역의 화질이 개선된 재생 영상을 나타낸다. 본 발명에 따르면, 게이즈 디텍션 방법을 이용하여 디스플레이되는 현재 멀티 프레임 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지한 후, 관심 영역을 포함하는 프레임 영상(510)만을 스케일러블 디코딩을 수행하여 관심 영역의 화질을 향상시키고 다른 프레임 영상(530)은 베이스 레이어 데이터만 디코딩한다.

66> II. 제2 실시예

67> 도 12는 본 발명의 제2 실시예에 따른 비디오 데이터 처리 장치의 블록도이다. 도 12를 참조하면, 비디오 데이터 처리 장치는 관심영역 결정부(710), 제어부(730), 데이터 통신부(750) 및 디코더(770)를 포함한다.

<68> 본 발명의 제2 실시예에 따르면, 전술한 바와 같은 게이지 디텍션 방법을 이용하여 관심 영역 결정부(710)가 디스플레이 되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지한다. 제어부(730)는 데이터 통신부(750)를 제어하여 관심영역 결정부(710)가 탐지한 관심 영역의 위치 정보를 본 발명의 제2 실시예에 따른 비디오 데이터 처리 장치로 비트스트림을 전송하는 상대방 장치(encoder)(도시되지 않음)로, 탐지된 관심 영역의 위치 정보를 전송한다.

<69> 탐지된 관심 영역의 위치 정보를 수신한 그 상대방 장치는 관심 영역을 포함하는 영상 객체만을 스케일러블 인코딩하고 다른 영상 객체들은 베이스 레이어 인코딩만 수행함으로써 통신 채널을 통해 전송할 데이터 양을 크게 줄인다. 즉, 데이터 전송 채널의 대역 폭의 한계를 고려하여 관심 영역의 화질을 개선한다.

<70> 데이터 통신부(750)를 통해 수신한 비트스트림은 디코더(770)로 입력된다. 디코더(770)는 제어부(730)의 제어에 따라 입력 비트스트림에 대해 스케일러블 디코딩을 수행한다. 전술한 제1 실시예에 설명된 디코더(150)와 같이, 입력 비트스트림에서 사용자가 응시하는 관심 영역을 포함하는 영상 객체와 나머지 영상 객체의 인핸스먼트 레이어 비트스트림을 선별할 필요가 없다. 상대방 장치에 의해 관심 영역을 포함하는 영상 객체만이 스케일러블 코딩되어 입력 비트스트림에는 관심 영역을 포함하는 영상 객체의 인핸스먼트 레이어 비트스트림만 포함되어 있기 때문이다.

<71> 한편, 데이터 통신 채널로는 PSTN, ISDN, Internet, ATM Network 또는 무선 통신망(Wireless communication network) 등과 같은 다양한 전송 매체가 이용될 수 있다. 데이터 통신 채널의 데이터 전송 속도가 떨어지는 경우, 상대방 장치에서 인코딩 시, 양자화 계수의 값을 증가시키는 등의 방법을 이용하여 베이스 레이어 데이터를 보다 열화시킴으로써 전송 데이터 양을 감소시킬 수도 있다.

<72> 이상, 본 발명의 바람직한 두 가지 실시예를 설명하였다. 본 발명에 따른 데이터 처리 장치는 양방향 영상 통신 시스템, 단방향 영상 통신 시스템 또는 다중 양방향 영상 통신 시스템에 적용할 수 있다.

<73> 양방향 영상 통신 시스템의 예로는 양방향 화상 회의(Video Teleconferencing) 또는 양방향 방송 시스템 등이 있으며, 단방향 영상 통신 시스템의 예로는 홈 쇼핑 방송과 같은 단방향 인터넷 방송 또는 주차장 감시 등의 감시 시스템(Surveillance system) 등이 있다. 끝으로 다중 양방향 영상 통신 시스템의 예로는 여러 사람 간의 화상 회의 시스템이 있다. 다만, 본 발명의 두 번째 실시예는 양방향 통신이 전제가 되므로 단방향 영상 통신에는 적용할 수 없다.

<74> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는, ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광데이터 저장 장치 등이 있으며, 또한 캐리어 웨이브(예를 들어, 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드로 저장되고 실행될 수 있다.

<75> 이제까지 본 발명에 대하여 몇 가지 바람직한 실시예들을 중심으로 설명하였으나, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 한다.

【발명의 효과】

- <76> 이상에서 설명한 바와 같이, 전송할 비디오 데이터 양이 매우 많으며, 데이터 전송 채널의 대역 폭의 제한 또는 디코더의 성능의 한계가 존재하여, 높은 화질의 영상을 사용자에게 재생하기 어려운 경우, 본 발명에 의하면, 게이즈 디텍션 방법을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 탐지하여 관심 영역의 화질만을 스케일러블 디코딩하여 향상시킴으로써 디코더의 로드를 감소시키거나 데이터 통신 채널의 대역 폭의 한계를 극복할 수 있다.

【특허청구범위】**【청구항 1】**

비디오 데이터 처리 방법에 있어서,

게이즈 디텍션을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계;

입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하는 단계; 및

상기 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 2】

제1 항에 있어서,

상기 입력 비트스트림은 다수의 영상 객체별로 스케일러블 코딩된 스케일러블 비트스트림인 것을 특징으로 하는 방법.

【청구항 3】

제1 항에 있어서,

상기 게이즈 디텍션은 상기 사용자의 머리의 움직임 또는 눈동자의 움직임을 추정하여 상기 관심 영역의 위치를 결정하는 것을 특징으로 하는 방법.

【청구항 4】

제2 항에 있어서,

상기 입력 비트스트림은 각 영상에 포함된 다수의 영상 객체들의 위치 정보를 포함하고

상기 비트스트림 선택 단계는 상기 관심 영역의 위치 정보와 상기 입력 비트스트림에 포함된 다수의 영상 객체들의 위치 정보를 비교하여, 상기 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림을 선택하는 것을 특징으로 하는 방법.

【청구항 5】

제2 항에 있어서,

상기 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체를 제외한 나머지 영상 객체들의 인헨스먼트 레이어 비트스트림을 선택하는 단계; 및

선택된 상기 나머지 영상 객체들의 인헨스먼트 레이어 비트스트림이 디코딩되지 않도록 폐기하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 6】

제1 항에 있어서,

상기 영상 객체는 멀티프레임 영상인 경우, 하나의 프레임 또는 하나의 프레임 영상을 다수의 영상 콘텐츠별로 분리하여 코딩한 경우, 영상 콘텐츠인 것을 특징으로 하는 방법.

【청구항 7】

비디오 데이터 처리 장치에 있어서,

입력 비트스트림을 스케일러블 디코딩하는 스케일러블 디코더;

게이즈 디텍션을 이용하여 디스플레이 되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하여, 상기 관심 영역의 위치 정보를 출력하는 관심영역 결정부; 및

상기 관심영역 결정부로부터 입력받은 상기 관심 영역의 위치 정보에 따라, 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림을 선택하고, 상기 선택된 베이스 레이어 비트스트림 및 인헨스먼트 레이어 비트스트림을 스케일러블 디코딩하도록 상기 스케일러블 디코더를 제어하는 제어부를 포함하는 것을 특징으로 하는 장치.

【청구항 8】

제7 항에 있어서,

상기 입력 비트스트림은 다수의 영상 객체별로 스케일러블 코딩된 스케일러블 비트스트림인 것을 특징으로 하는 장치.

【청구항 9】

제7 항에 있어서,

상기 게이즈 디텍션은 상기 사용자의 머리의 움직임 또는 눈동자의 움직임을 추정하여 상기 관심 영역의 위치를 결정하는 것을 특징으로 하는 장치.

【청구항 10】

제8 항에 있어서,

상기 입력 비트스트림은 각 영상에 포함된 다수의 영상 객체들의 위치 정보를 포함하고

상기 제어부는 상기 관심 영역의 위치 정보와 상기 입력 비트스트림에 포함된 다수의 영상 객체들의 위치 정보를 비교하여, 상기 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하는 것을 특징으로 하는 장치.

【청구항 11】

제8 항에 있어서,

상기 제어부는 상기 입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체를 제외한 나머지 영상 객체들의 인핸스먼트 레이어 비트스트림을 선택하고, 상기 나머지 영상 객체들의 인핸스먼트 레이어 비트스트림이 디코딩하지 않도록 상기 스케일러블 디코더를 제어하는 것을 특징으로 하는 장치.

【청구항 12】

제7 항에 있어서,

상기 영상 객체는 멀티프레임 영상인 경우, 하나의 프레임 또는 하나의 프레임 영상을 다수의 영상 콘텐츠별로 분리하여 코딩한 경우, 영상 콘텐츠인 것을 특징으로 하는 장치.

【청구항 13】

비디오 데이터 처리 방법에 있어서,

상대방 장치로부터 수신한 이전 비트스트림을 디코딩하여 사용자에게 디스플레이하는 단계;

게이즈 디텍션을 이용하여 디스플레이되는 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계;

상기 관심 영역의 위치 정보를 상기 상대방 장치로 송신하는 단계;

상기 상대방 장치로부터 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 포함하는 현재 비트스트림을 수신하는 단계; 및

상기 현재 비트스트림을 스케일러블 디코딩하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 14】

제13 항에 있어서,

상기 현재 비트스트림은 하나의 영상에 포함된 다수의 영상 객체 중 상기 관심 영역을 포함하는 영상 객체만 스케일러블 코딩된 비트스트림인 것을 특징으로 하는 방법.

【청구항 15】

제13 항에 있어서,

상기 게이지 디텍션은 상기 사용자의 머리의 움직임 또는 눈동자의 움직임을 추정하여 상기 관심 영역의 위치를 결정하는 것을 특징으로 하는 방법.

【청구항 16】

제13 항에 있어서,

상기 영상 객체는 멀티프레임 영상인 경우, 하나의 프레임 또는 하나의 프레임 영상을 다수의 영상 콘텐츠별로 분리하여 코딩한 경우, 영상 콘텐츠인 것을 특징으로 하는 방법.

【청구항 17】

비디오 데이터 처리 장치에 있어서,

입력 비트스트림을 스케일러블 디코딩하는 스케일러블 디코더;

상대방 장치로부터 수신하여 디코딩된 후 사용자에게 디스플레이 되는 영상에서 상기 사용자가 응시하는 관심 영역의 위치를 게이지 디텍션을 이용하여 결정하여, 상기 관심 영역의 위치 정보를 출력하는 관심영역 결정부; 및

상기 관심 영역의 위치 정보를 상기 상대방 장치로 송신하는 데이터 통신부를 포함하고

상기 스케일러블 디코더는 상기 상대방 장치로부터 수신한, 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 포함하는 현재 비트스트림을 스케일러블 디코딩하는 것을 특징으로 하는 장치.

【청구항 18】

제17 항에 있어서,

상기 현재 비트스트림은 하나의 영상에 포함된 다수의 영상 객체 중 상기 관심 영역을 포함하는 영상 객체만 스케일러블 코딩된 비트스트림인 것을 특징으로 하는 장치.

【청구항 19】

제17 항에 있어서,

상기 게이지 디텍션은 상기 사용자의 머리의 움직임 또는 눈동자의 움직임을 추정하여 상기 관심 영역의 위치를 결정하는 것을 특징으로 하는 장치.

【청구항 20】

제17 항에 있어서,

상기 영상 객체는 멀티프레임 영상인 경우, 하나의 프레임 또는 하나의 프레임 영상을 다수의 영상 콘텐츠별로 분리하여 코딩한 경우, 영상 콘텐츠인 것을 특징으로 하는 장치.

【청구항 21】

게이즈 디택션을 이용하여 디스플레이되는 현재 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계;

입력 비트스트림에서 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 선택하는 단계; 및

상기 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 스케일러블 디코딩하는 단계를 포함하는 것을 특징으로 하는 비디오 데이터 처리 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

【청구항 22】

상대방 장치로부터 수신한 이전 비트스트림을 디코딩하여 사용자에게 디스플레이하는 단계;

게이즈 디택션을 이용하여 디스플레이되는 영상에서 사용자가 응시하는 관심 영역의 위치를 결정하는 단계;

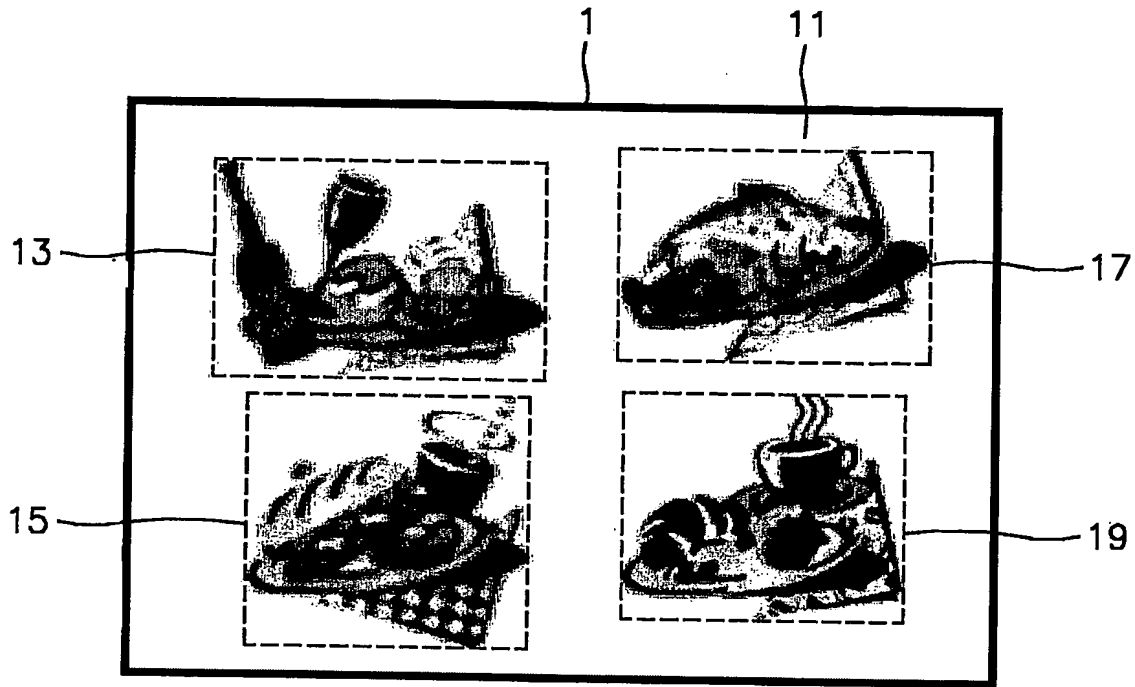
상기 관심 영역의 위치 정보를 상기 상대방 장치로 송신하는 단계;

상기 상대방 장치로부터 상기 관심 영역을 포함하는 영상 객체의 베이스 레이어 비트스트림 및 인핸스먼트 레이어 비트스트림을 포함하는 현재 비트스트림을 수신하는 단계; 및

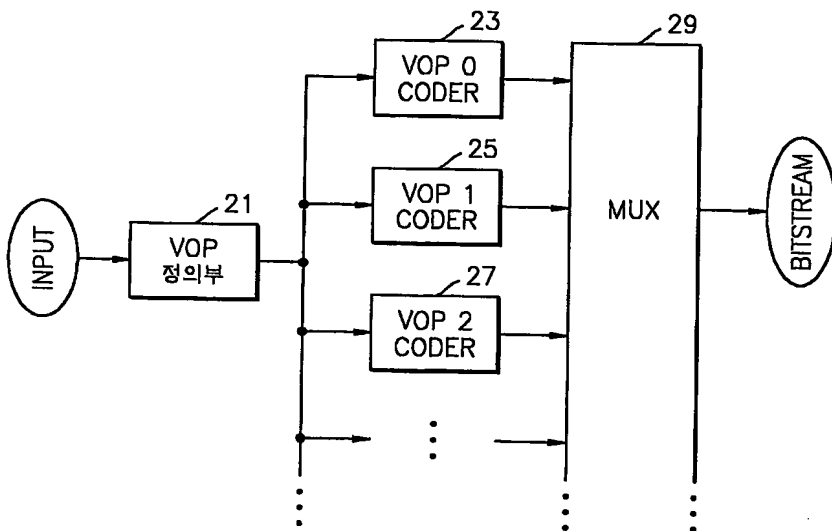
상기 현재 비트스트림을 스케일러블 디코딩하는 단계를 포함하는 것을 특징으로 하는 비디오 데이터 처리 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

【도면】

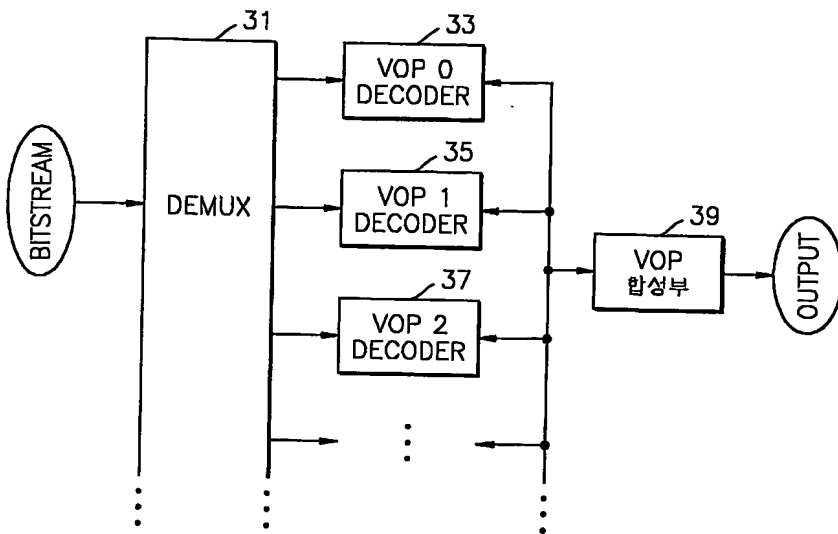
【도 1】



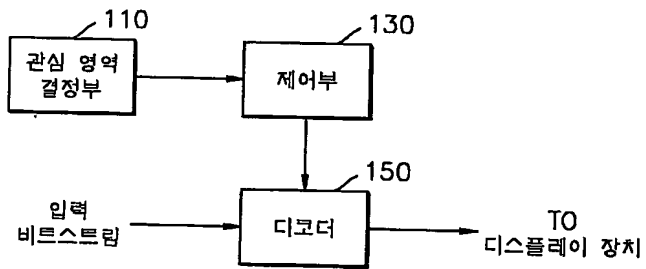
【도 2】



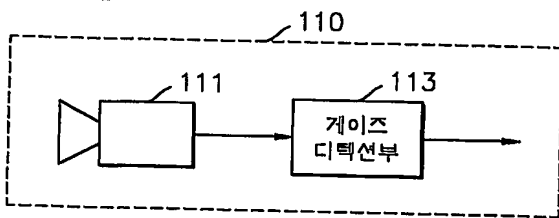
【도 3】



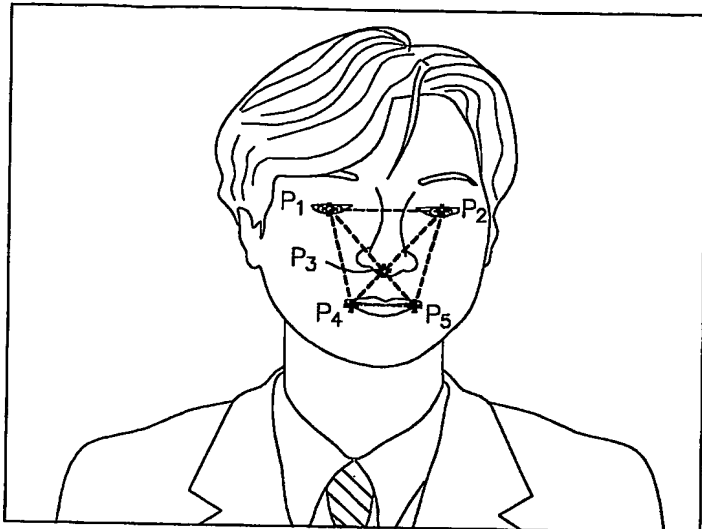
【도 4】



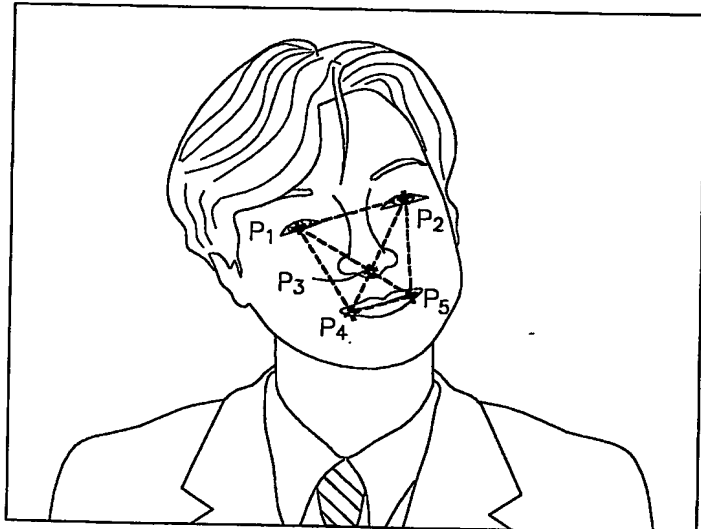
【도 5】



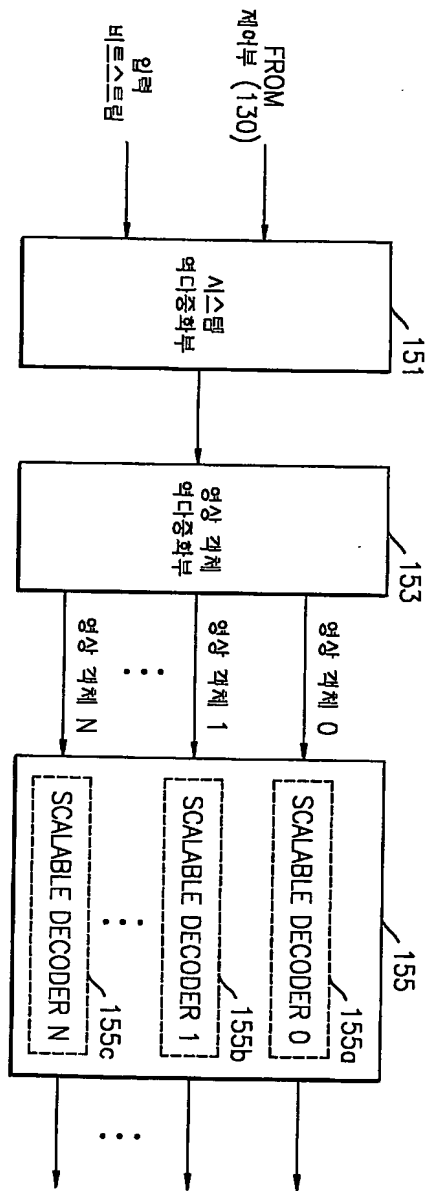
【도 6a】



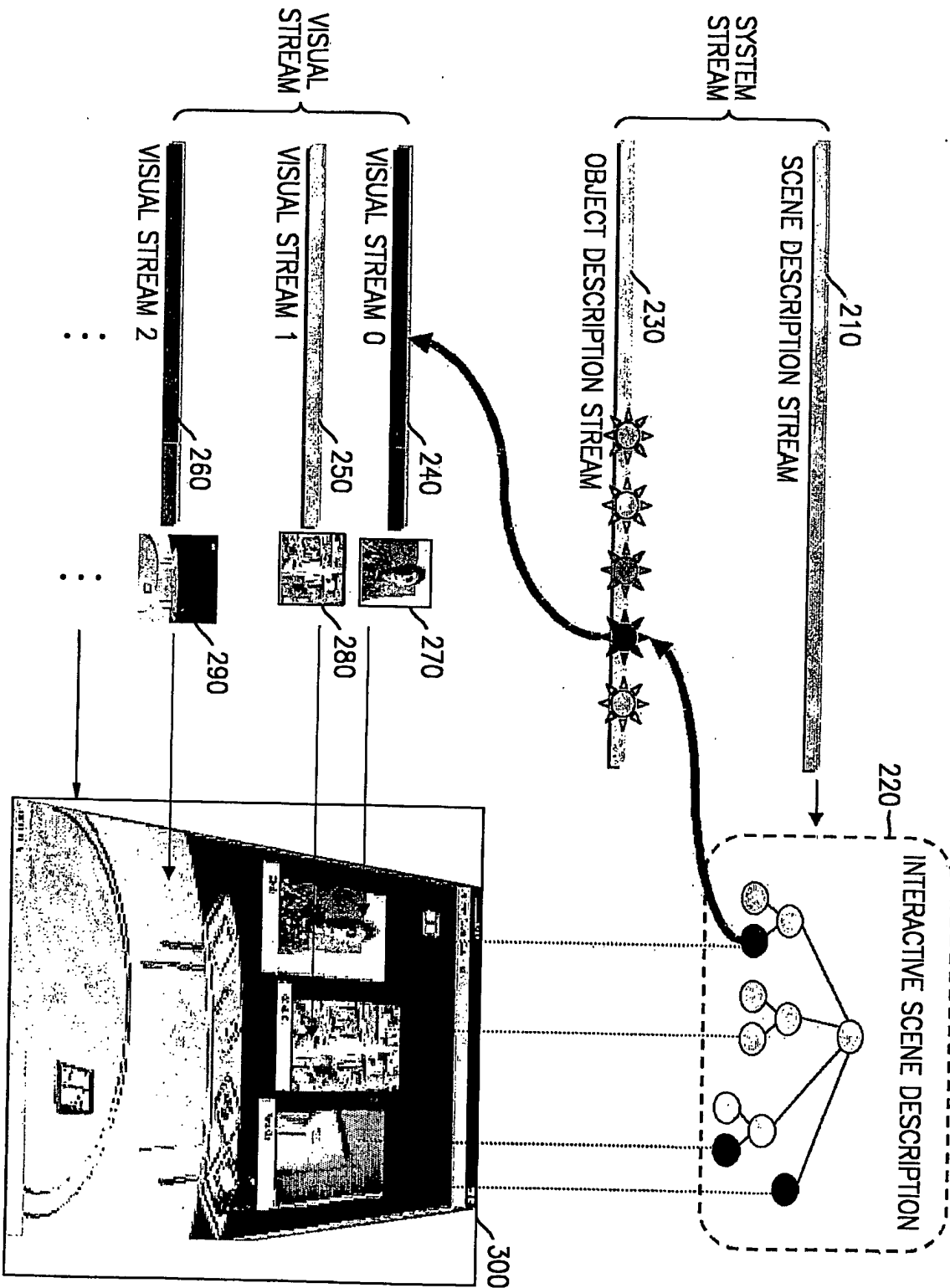
【도 6b】



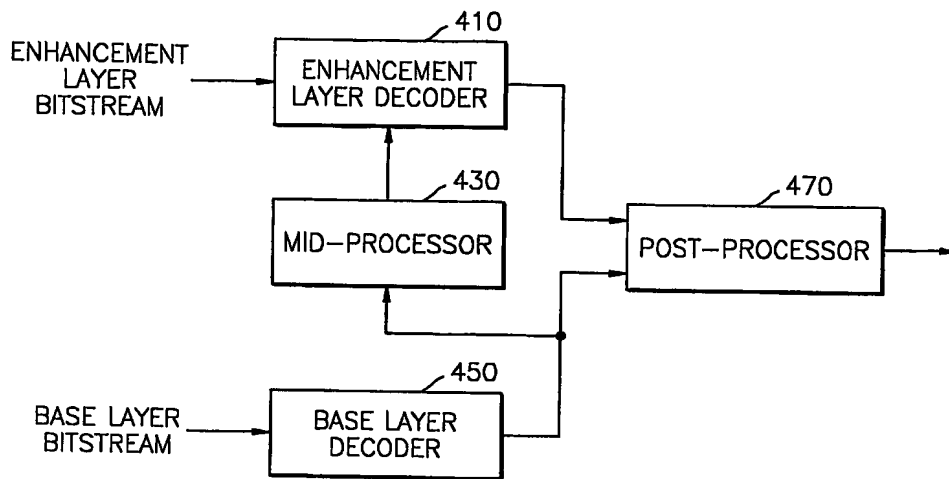
【도 7】



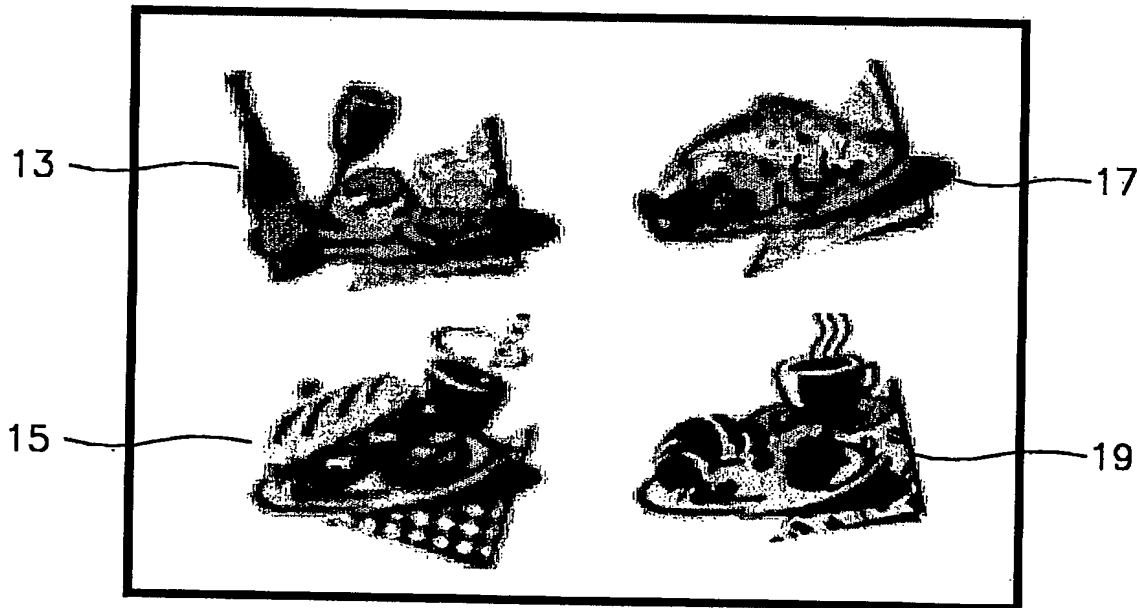
【도 8】



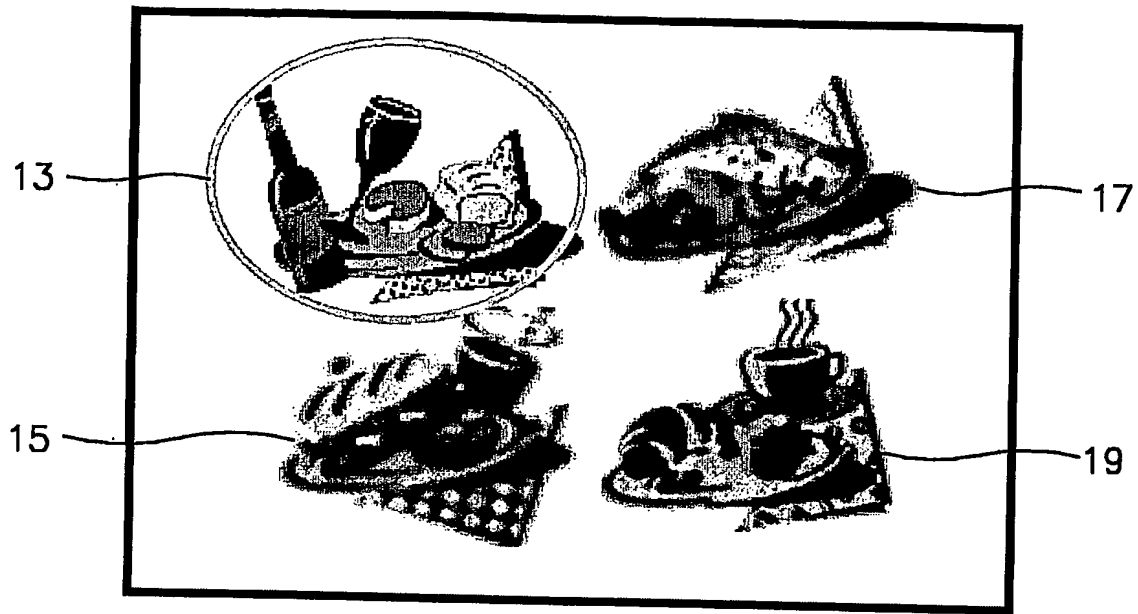
【도 9】



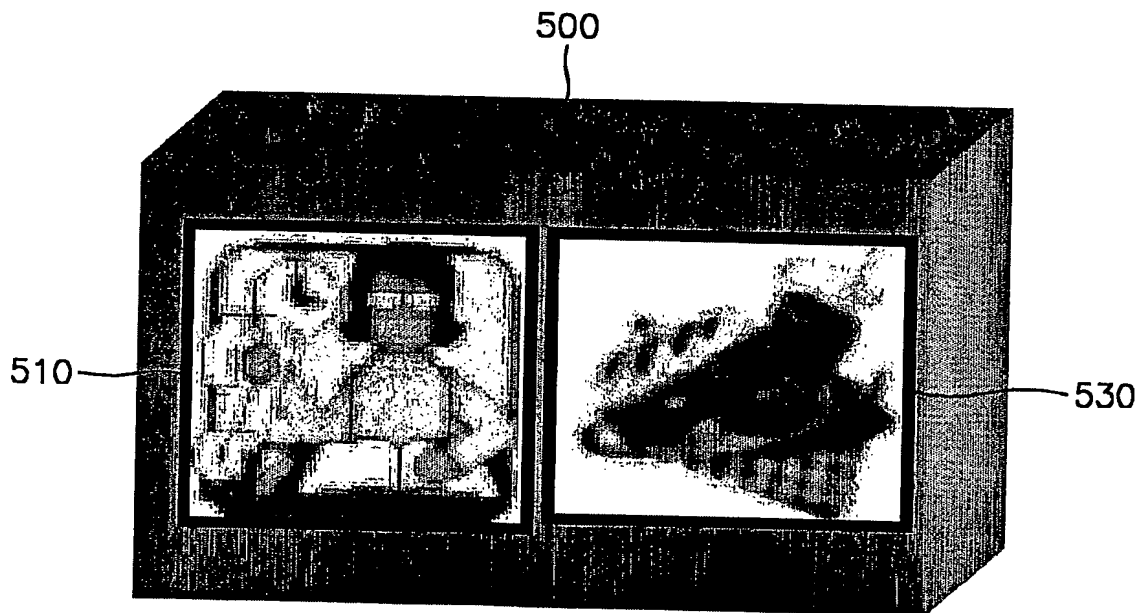
【도 10a】



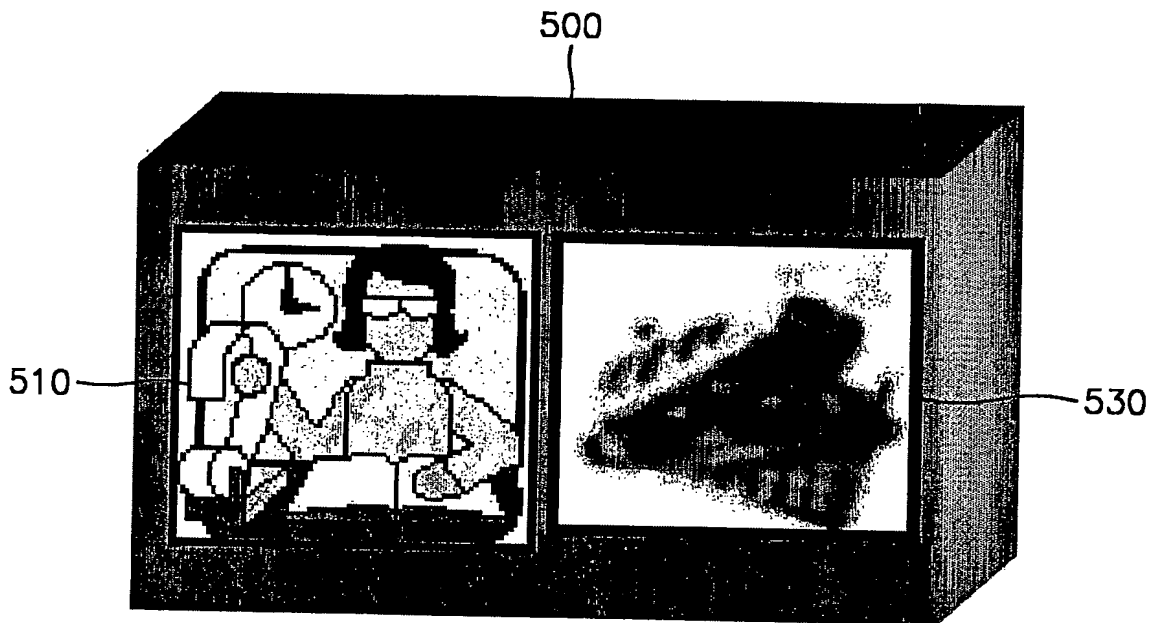
【도 10b】



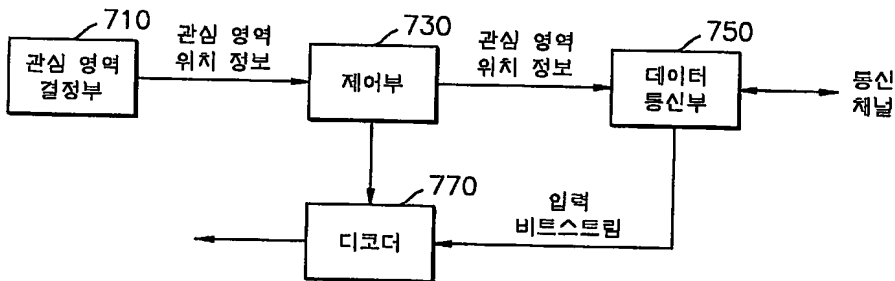
【도 11a】



【도 11b】



【도 12】



This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**